International Journal of Engineering Practical Research
IJEPR Volume 1, Issue 1, August 2012 PP. 21-25

Use of Complex Discrete Signals for Steganographic Information Security

Alexandr Kuznetsov, Alexey Smirnov
Kozhedub Air Force University, Kharkov, Ukraine
Department of Software, Kirovohrad National Technical University, Kirovohrad, Ukraine
kuznetsov_alex@rambler.ru, assa_s@mail.ru

(Abstract) The steganographic methods of information embedding in stable images for the secure transfer of information are investigated. The method of steganographic security based on the use of complex discrete signals and technology of direct expansion of spectrum is examined, its efficiency is probed from point of the provided security, carrier capacity and amount of introduced distortions in a container-image.

Keywords: Container-image; Steganography; Noise-immunity; Erroneous Extraction.

1. RAISING THE PROBLEM IN A GENERAL VIEW AND ANALYSIS OF LITERATURE

Steganographic systems are an important direction in the development of modern facilities of information security. They secure from the opponent not only informative content of transferred information but the fact of passing the messages itself [1, 2]. Steganographic method based on the developed mathematical apparatus of discrete signals theory and noise-immune communication [3 - 21].

This research studies the steganographic method of information protection in stable images [2, 10-11, 20], based on the use of complex discrete signals and technology of direct expansion of a spectrum, estimation of its efficiency from the point of provided security, carrier capacity and amount of the insertion distortions in a container-image.

2. DIRECT EXPANSION OF SPECTRUM (DIRECT SEQUENCE) IN A COMMUNICATION THEORY

The methods of discrete signals theory, cross-correlation and spectral analysis are utilized [3 – 5, 21] for the construction of modern noise-immune digital communication networks. Broadband systems with complex discrete signals and direct expansion of spectrum [3, 4] are considered to be the most perspective from the point of the effective use of frequency-time and energy resources of communication channels.

We will consider an informative signal, which is a set of separate values taken at certain moments of time, as a discrete signal. We will further consider a complex discrete signal as a binary pseudorandom sequence (BPRS) $\Phi_i = \left(\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}}\right)$ of length of n from set $\Phi = \left\{\Phi_0, \Phi_1, ..., \Phi_{M-1}\right\}$ with power of $|\Phi| = M$. The elements of binary BPRS take one of the following values:

$$\varphi_{i_z} = \begin{cases} +1 \\ -1 \end{cases}, \quad z = 0,...,n-1 \tag{1}$$

For the construction of noise-immune broadband connection the concept of correlation of discrete signals to statistical relationship of two or a few BPRS is utilized. The mathematical measure of correlation (conformity) of two discrete signals $\Phi_i, \Phi_j \in \Phi$ is the coefficient of correlation $\rho(\Phi_i, \Phi_j)$ [3, 4]:

$$\rho\left(\Phi_i, \Phi_j\right) = \frac{1}{n} \sum_{z=0}^{n-1} \Phi_{i_z} \Phi_{j_z} \tag{2}$$

Two signals Φ_i, Φ_j are orthogonal, if coefficient of correlation $\rho(\Phi_i, \Phi_j) = 0$. If $\rho(\Phi_i, \Phi_j) \approx 0$, we will name signals Φ_i and Φ_j quasi-orthogonal[5, 7].

In a modern digital communication theory the large assemblies of weak-correlated discrete signals are utilized for the construction of the wide-band noise-immune systems of data communication. Transferred messages in such channels acquire the appearance of noise-like sequences, and due to high power of assemblies of discrete signals and direct expansion of frequency spectrum, high authentication, noise-immunity and secrecy of the digital connection channels are provided [3 – 5].

For communication of data in a broadband communication network, informative signal:

$$x(t) = \begin{cases} +1 \\ -1 \end{cases},$$

is modulated by means of multiplying by an extending code signal $g(t) = \Phi_i \in \Phi$ – pseudorandom sequence of the assemblies of discrete signals which were considered earlier. As code signal due to its statistical properties is similar to noise, therefore the extended signal:

$$y'(t) = y(t) + e(t) \tag{3}$$

is poorly distinguished from noises in a communication channel, which allows to carry out a hidden transmission.

During the reception in a demodulator the received signal $y'(t) = y(t) + e(t)$, as a mixture of the transferred sequence $y(t)$ and errors in a communication channel $e(t)$, is multiplied by the synchronized copy of extending signal $g(t)$. In other words, on a receiving side the calculation of coefficient of correlation is carried out (2). the value of such coefficient is determined by a decision-making rule:

$$\rho\left(y'(t), g(t)\right) = \frac{1}{n}\sum_{z=0}^{n-1} x(t)\Phi_{i_z}\Phi_{i_z} + \frac{1}{n}\sum_{z=0}^{n-1} e(t)\Phi_{i_z}$$

Taking into account a pseudo randomness of Φ_i used as $g(t)$, the second term can be neglected (the amount of «+1» is approximately equal to the amount «-1») in the right part of equation, i.e.

$$\rho\left(y'(t), g(t)\right) \approx \rho\left(y(t), g(t)\right) = x(t)\frac{1}{n}\sum_{z=0}^{n-1}\left(\Phi_{i_z}\right)^2 = x(t), \quad (4)$$

i.e. the value of informative signal on a receiving side is determined by the following expression:

$$x(t) = \begin{cases} +1, \ for \ \rho\left(y'(t), g(t)\right) \approx +1; \\ -1, \ for \ \rho\left(y'(t), g(t)\right) \approx -1; \end{cases} \quad (5)$$

where the sign « ≈ » is imposed by the presence of errors $e(t)$, caused by natural or intentional noise in a communication channel.

3. DIRECT EXPANSION OF SPECTRUM IN STEGANOGRAPHY

In the method of Smith-Comiskey [2], as well as in the considered above communication networks with direct expansion of spectrum, an informative message is bitwise modulated by multiplying the assembly of orthogonal signals. Then a modulated message is embedded into a container – stable image.

We will present an informative message m which is built in a digital container-image as blocks m_i of equal length, i.e. $m = (m_0, m_1, ..., m_{N-1})$, where every block m_i is a sequence (vector) of n bits: $m_i = (m_{i_0}, m_{i_1}, ..., m_{i_{n-1}})$. a container-image will be considered as an array of data C by a dimension of $K \cdot L$, divided onto sub-blocks with a length of $k \cdot l = n$. In the capacity of array component C, for example, raster information of in-use image may be taken. Confidential key data is a set of base functions $Key = \Phi = \{\Phi_0, \Phi_1, ..., \Phi_{M-1}\}$, where all the base functions $\Phi_i = (\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}})$ are mutually orthogonal discrete signals with the length equal to the size n block of report m_i, i.e. for any $i, j \in [0, ..., M-1]$ the following equality is true:

$$\rho\left(\Phi_i, \Phi_j\right) = \frac{1}{n}\sum_{z=0}^{n-1}\Phi_{i_z}\Phi_{j_z} = \begin{cases} +1, \text{при i = j;} \\ -1, \text{при } i \neq \text{j;} \end{cases}$$

The purpose of the steganographic transformation of information is the embedding of every separate block of message m_i in a proper block of a container-image. In the block of data of digital representation a dimension of $K \cdot L$ elements/

$K \cdot \frac{L}{n}$ blocks of informative message can be embedded, i.e. to $K \cdot L$ bits. we will utilize assemblies of orthogonal Walsh-Hadamard discrete signals as key data (an array of base functions $Key = \Phi$).

Embedding of informative message is carried out as follows. Every block of a message $m_{i_j}, j = 0, ..., n-1$ is compared to a separate block of a container-image. Every informative bit of a block $m_{i_j}, j = 0, ..., n-1$ appears as an informative signal:

$$m_{i_j}(t) = \begin{cases} +1, m_{i_j} = 1; \\ -1, m_{i_j} = 0; \end{cases}$$

and by analogy with (3) is modulated by extending code signal (by base functions), i.e. BPRS $\Phi_j \in \Phi$. As a result, for every informative block m_i the modulated informative signal is formed:

$$E_i(t) = \sum_{j=0}^{n-1}\sum_{z=0}^{n-1} m_{i_j}(t)\Phi_{j_z}. \quad (6)$$

A derived block of a message E_i is pixel-wise added to a sub-block C_i of a container, and a steganogram (filled container) is formed by means of aggregation of data arrays $S_i, i = 0, ..., N-1$ as follows:

$$S_i = C_i + E_i \cdot G, \quad (7)$$

where $G > 0$ is an amplification of extending signal factor, defining the «energy» of the embedded bits of informative sequence.

The operation of decoding is the renewal of a hidden message by projection of every block S_i, of received steganographic image S on all of the basis functions $\Phi_j \in \Phi$, $i = 0, ..., N-1$. For this purpose, every block S_i is presented in the form a vector $S_i = (S_{i_0}, S_{i_1}, ..., S_{i_{n-1}})$, $i = 0, ..., N-1$. To extract j-th bit of a message from i-th block of steganographic image, it is necessary to calculate the coefficient of correlation between Φ_j and the accepted block S_i (presented as a vector):

$$\rho\left(S_i, \Phi_j\right) = \frac{1}{n}\sum_{z=0}^{n-1} S_{i_z}\Phi_{j_z} = G \cdot \frac{1}{n}\sum_{z=0}^{n-1} E_{i_z}\Phi_{j_z} + \frac{1}{n}\sum_{z=0}^{n-1} C_{i_z}\Phi_{j_z}, \quad (8)$$

where C_i is one-dimensional array, i.e. the proper block of a container, presented in the form of a vector.

An array C_i is assumed to have a random statistical structure, i.e. the second element in the right part of the expression (8) is close to zero and can be ignored. Then we have:

$$\rho(S_i,\Phi_j)\approx G\cdot E_i\cdot\Phi_j=G\cdot\sum_{l=0}^{n-1}\sum_{z=0}^{n-1}m_{i_X}(t)\cdot\Phi_{l_z}\Phi_{j_z}. \quad (9)$$

By analogy with (6) we notice that all the sequences from a set of Φ are mutually ortogonal, i.e. if $l\neq j$, we have $\rho(\Phi_i,\Phi_j)=0$. Thus, all the terms in the right part of (9) in case of $l\neq j$ can be ignored. Thus:

$$\rho(S_i,\Phi_j)\approx G\cdot m_{i_j}(t)\cdot\frac{1}{n}\sum_{z=0}^{n-1}(\Phi_{j_z})^2=G\cdot m_{i_j}(t). \quad (10)$$

Since $G>0$ and $n>0$, a sign of $\rho(S_i,\Phi_j)$ in (10) depends only on $m_{i_j}(t)$, whence it appears:

$$m_{i_j}(t)=sign(\rho(S_i,\Phi_j))=\begin{cases}-1,\ for\ \rho(S_i,\Phi_j)<0;\\ +1,\ for\ \rho(S_i,\Phi_j)>0;\\ ?,\ for\ \rho(S_i,\Phi_j)=0;\end{cases} \quad (11)$$

If $\rho(S_i,\Phi_j)=0$ in (11) then we assume, that the built-in information was lost.

The considered steganographic system inherits all of advantages of broadband communication networks: resistance to unauthorized extraction of built-in reports and its destruction or modification.

4. ESTIMATION OF STEGANOGRAPHIC SYSTEM EFFICIENCY

Taking into account the purpose of steganographic system, we will introduce the following efficiency indexes.

1. *Throughput* – the ratio of V embedded in a container of information to the general capacity D of a container

$$Q=\frac{V}{D}. \quad (12)$$

2. *Key data capacity* (bits)

$$l_{Key}=\log_2(|Key|), \quad (13)$$

where $|Key|$ is the power of a set of key data.

3. *Security of the steganographic method* we will estimate as an inverse value to the power of a set of secret key data. We can consider it as a probability index of a secret key selection:

$$W=\frac{1}{|Key|}=2^{-l_{Key}}. \quad (14)$$

4. *Value of added distortions* is a percentage ratio of an arithmetical mean value of all the absolute values Δ-changes of a container data to a maximum value $\Delta_{\max}$:

$$I=\frac{\Delta_{mean}}{\Delta_{\max}}\cdot 100=\frac{100}{\Delta_{\max}\cdot D}\cdot\sum_{i=1}^{D}|\Delta_i|, \quad (15)$$

where Δ_i – Δ-changes i-th element of a container.

5. *Probability of erroneous extraction* of message information data:

$$P_{err}=\lim_{D\to\infty}\frac{V_{err}}{D}=1-\lim_{D\to\infty}\frac{V-V_{err}}{D}, \quad (16)$$

where V_{err} is the capacity of extracted data.

By using of introduced indexes (12) – (16) we will estimate the efficiency of the considered steganographic method of information security. Experimental researches will be conducted by embedding of information in bitmapped images (color model of RGB) with an 8-bit encoding of every color. The achieved results are presented in **figures** 1 – 4.

From the presented in a **fig. 1** dependences it follows, that the increase of the carrying capacity of a steganographic channel conduces to the strong increase of introduced distortions in a container-image. Unnoticeable for an observer distortions (that are lying below than threshold of sensitiveness of the visual system of a human) are brought in only at $Q\le 0.005$. It corresponds to building not more than 10 bits in one block of an image, i.e. to modulating up to ten informative signals $m_{i_j}(t)$, $j=0,...,9$ in the expression (6). Dependences, presented on **fig. 2** and **3** testify that an amplification factor used in expressions (7) - (10) allows substantially reduce the probability of erroneous extraction of informative data. Unfortunately, it is achieved at the expense of a strong increase of the introduced distortions in a utilized container-image. The dependences are achieved at $Q=0.005$. It is obvious, that for such a value of carrying capacity, an amplification factor cannot exceed 1..1,5 (see **fig. 2**).

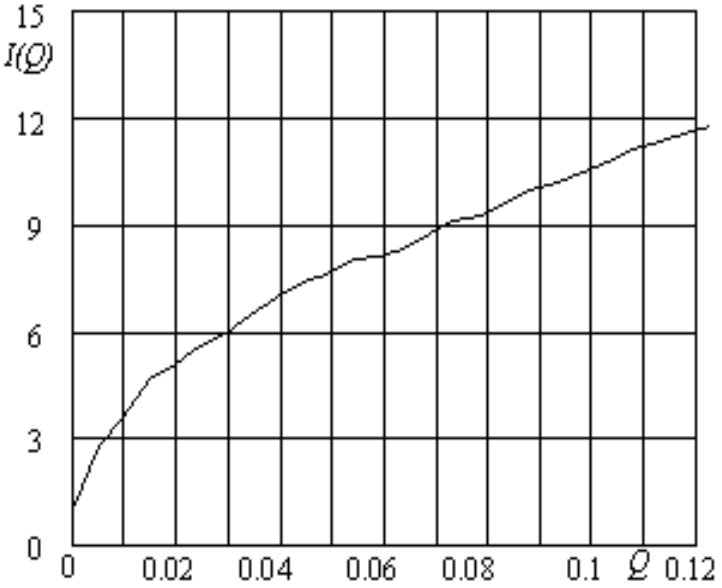

Figure 1. Dependence $I(Q)$

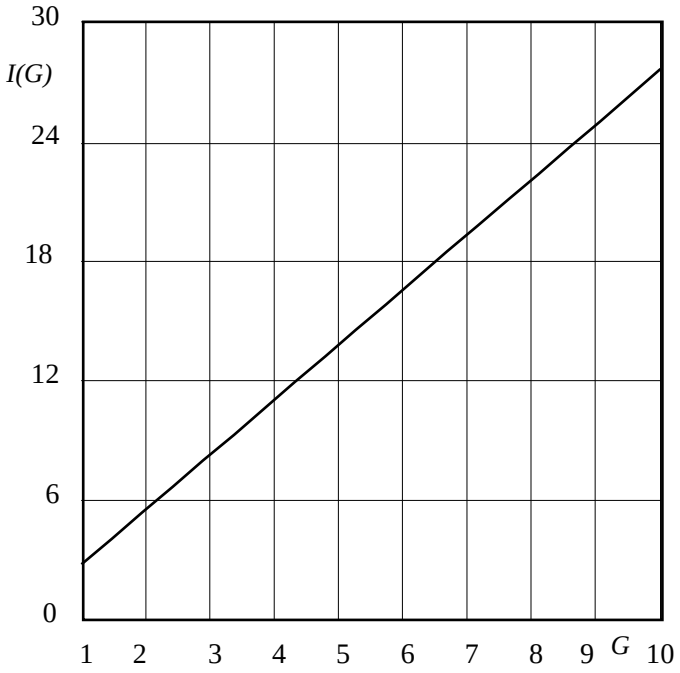

Figure 2. Dependence $I(G)$ for $Q = 0.005$

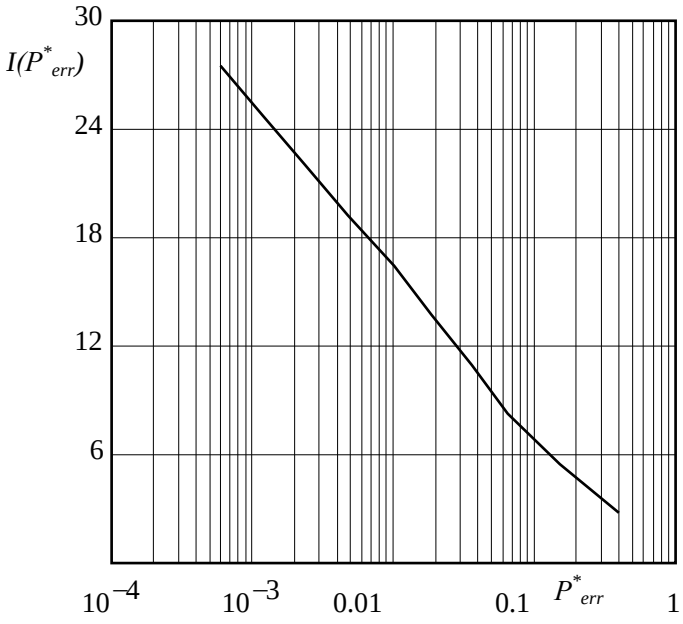

Figure 3. Dependence $P^*_{err}(G)$ for $Q = 0.005$

However, even for such values the probability of an erroneous extraction is strong and lies in a range of 0,1 .. 0,5. Integral dependence $I(P^*_{err})$, presented on a **fig. 4**, summarizes the information on **fig. 2** and **3**. For a fixed carrying capacity $Q = 0.005$, an empiric curve, characterizing the dependence of size of the introduced distortions in a container-image and probabilities of erroneous extraction of informative data, is derived.

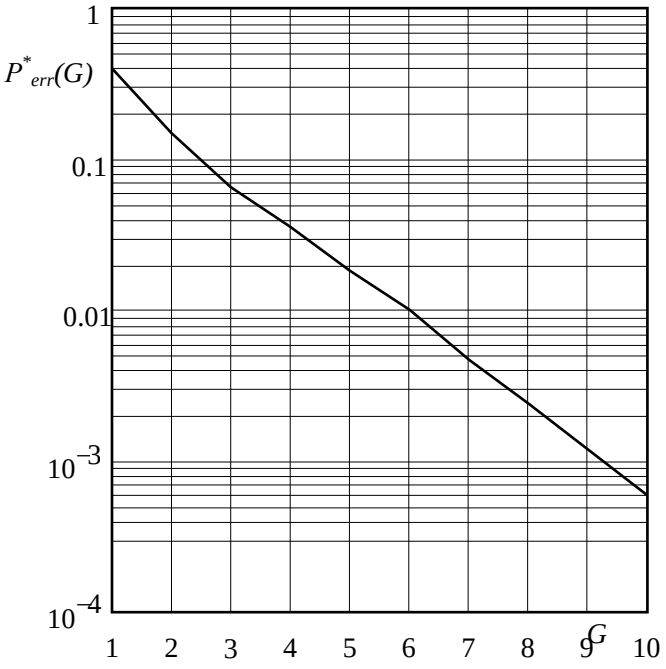

Figure 4. Dependence $I(P^*_{err})$ for $Q = 0.005$

Obtaining low distortions lying below the threshold of visual human sensitiveness ($I \le 2...3\%$) For $Q = 0.005$ is possible only at even higher probability of erroneous extraction of informative data ($P_{err} \ge 0.1$). It is obvious, that practical application of similar steganographic systems has to be combined with the anti-noise coding of informative data, that allows to substantially reduce P_{err}.

5. CONCLUSIONS

The conducted researches show that the use of direct expansion of discrete signals' spectrum for steganographic purposes allows carrying out secretive embedding of information messages in stable images. The task of an extraction of message on the receiving side of steganographic system is equivalent to the task of finding out the information from a mixture of desired signal and noise (distortions) in a broadband communication network.

During the researches the following limitations of the steganographic systems with the expansion of discrete signals' spectrum are discovered: the probability of correct extraction of embedded data depends on a size of the introduced distortions, which, in its turn, depends on the provided carrying capacity of a steganographic channel. In other words, the practical construction of a steganographic system involves the search of a compromise between the size of introduced distortions, the probability of correct extractions of a message on a receiving side, as well as the provided carrying capacity. In addition, during the researches it is revealed that the probability of correct extraction of

embedded data directly depends on statistical properties of in-use container-image.

The perspective direction of further researches, on the authors' opinion, is the use of large assemblies of weak-correlated (quasi-orthogonal) discrete signals for the construction of steganographic systems with direct expansion of a spectrum. It allows, on the one hand, increasing substantially the carrying capacity of a steganographic channel without the considerable increase of the introduced distortions in a container-image. On the other hand, it allows to substantially reduce a probability of erroneous extraction of built-in data by means of the adaptive forming (choice) of discrete signals on a criterion of minimization of the correlation coefficient with a container-image.

ACKNOWLEDGMENT

The authors thank Gorbenko I.D and Dolgov V.I. for their valuable comments, which helped to improve the paper.

REFERENCES

[1] Konakhovich G. F., Puzyrenko A. Y. Computer steganography. Theory and practice. – К.: «МК-Press», 2006. - 288 p., il.

[2] J. Smith, B. Comiskey, Modulation and Information hiding in Image. // Information hiding: First Int. Workshop "InfoHiding'96", Springer as Lecture Notes in Computing Science, vol 1174. 1996. – pp. 207-227.

[3] Digital methods in space communication. /Edited by S. Golomba.- M.: Sviaz, 1969. – 272 p.

[4] Skliar B. Digital communacations. Fundamentals and applications. – New Yersey: Prentice Hall PTR, 2001. – 1104 p.

[5] Arko Kundu, Kaushik Chakraborty and Souvik Bhattacharyya, Data Hiding in Images Using Mod 16 Method, in Inthe Proceedings of ETECE 2011, (2011).

[6] Souvik Bhattacharyya, Indradip Banerjee and Gautam Sanyal,A Survey of Steganography and Steganalysis Technique in Image, Text, Audio and Video as Cover Carrier,in Journal of Global Research in Computer Science(JGRCS) VOL 2, NO 4 (2011), APRIL-2011.

[7] Natarajan Meghanathan and Lopamudra Nayak, Steganalysis Algorithms for Detecting the Hidden Information in Image, Audio and Video Cover Media, in at International Journal of Network Security & Its Application (IJNSA),Vol.2, No.1, January 2010.

[8] Souvik Bhattacharyya and Gautam Sanyal, Implementa-tion and Design of an Image based Steganographic model,in Proceedings of IEEE International Advance ComputingConference, (2009).

[9] Souvik Bhattacharyya and Gautam Sanyal,An Imagebased Steganography model for promoting Global Cyber Security, in Proceedings of International Conference onSystemics, Cybernetics and Informatics, (2009).

[10] Marvel, L. M., C. G. Boncelet, Jr., and C. T. Retter. Spread Spectrum Image Steganography, IEEE Transactions on Image Processing, Vol 8, No 8, August 1999, pages 1075-1083.

[11] L. M. Marvel and C. T. Retter, "The Use of Side Information in Image Steganography," Proceedings of the IEEE 2000 International Symposium on Information Theory and Its Applications (ISITA'2000), Honolulu, HI, November 2000.

[12] Kruus, P., Scace, C., Heyman, M., Mundy, M.: A survey of steganographytechniques for image files, Advanced Security Research Journal, vol 5, no 1, p.41552 (2003).

[13] Levicky, D., Šurin, S.: Codebook code division multiple access image steganography, Radioengineering, vol. 14, no. 1, p. 41545l (2005).

[14] Stefan Hetzl, "A Survey of Steganography", January 2002.

[15] Tu Ran, "Steganography: The Art of Hiding Data", Spring 2002.

[16] Morkel, T., Eloff, J.H.P, Olivier, M.S.: An overview of image steganography, Proceeding of the Fifth Annual Information Security South Africa Conference(ISSA2005) (2005).

[17] Morkel, T., Eloff, J.H.P, Olivier, M.S.: An overview of image steganography, Proceeding of the Fifth Annual Information Security South Africa Conference(ISSA2005) (2005).

[18] Dunbar, B., "Steganographic techniques and their use in an Open-Systems environment", SANS Institute, January 2002.

[19] Venkatraman, S., Abraham, A. & Paprzycki, M., "Significance of Steganography on Data Security", Proceedings of the International Conference on Information Technology: Coding and Computing, 2004.

[20] Patent No.: US 6,557,103 B1, Int.Cl. G06F 11/30. Charles G. Boncelet, Jr., Lisa M. Marvel, Charles T. Retter. Spread Spectrum Image Steganography. Patent No.: US 6,557,103 B1, Int.Cl. G06F 11/30. – № 09/257,136; Filed Feb. 11, 1999; Date of Patent Apr. 29, 2003.

[21] Gorbenko I.D., Stasiev Y.V. Analysis of derived ortogonal signal systems // Radiotekhnika. – 1989. – № 9. – pp. 16 – 18.

Author Introduction

Alexandr Kuznetsov, Kharkov Kozhedub Air Force University, Kharkov, Ukraine.

The field of scientific interests: the coding interference immunity theory, methods of complex signals forming, cryptography and authentication.

Alexey Smirnov, Kirovohrad National Technical University, Kirovohrad, Ukraine.

The field of scientific interests: information security and routing issues.